

PCT/M

00/00473

07.08.00

4
KONINKRIJK DER



NEDERLANDEN

1100/473

Bureau voor de Industriële Eigendom



REC'D 23 AUG 2000

WIPO PCT

Hierbij wordt verklaard, dat in Nederland op 5 juli 1999 onder nummer 1012519,

ten name van:

PTT POST HOLDINGS B.V.

te Den Haag

een aanvraag om octrooi werd ingediend voor:

"Inrichting en werkwijze voor het actualiseren van een adressendatabase met geregistreerde adresrecords",

en dat de hieraan gehechte stukken overeenstemmen met de oorspronkelijk ingediende stukken.

Rijswijk, 7 augustus 2000.

De Directeur van het Bureau voor de Industriële Eigendom,
voor deze,

Mw. I.W. Scheevelenbos-de Reus.

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

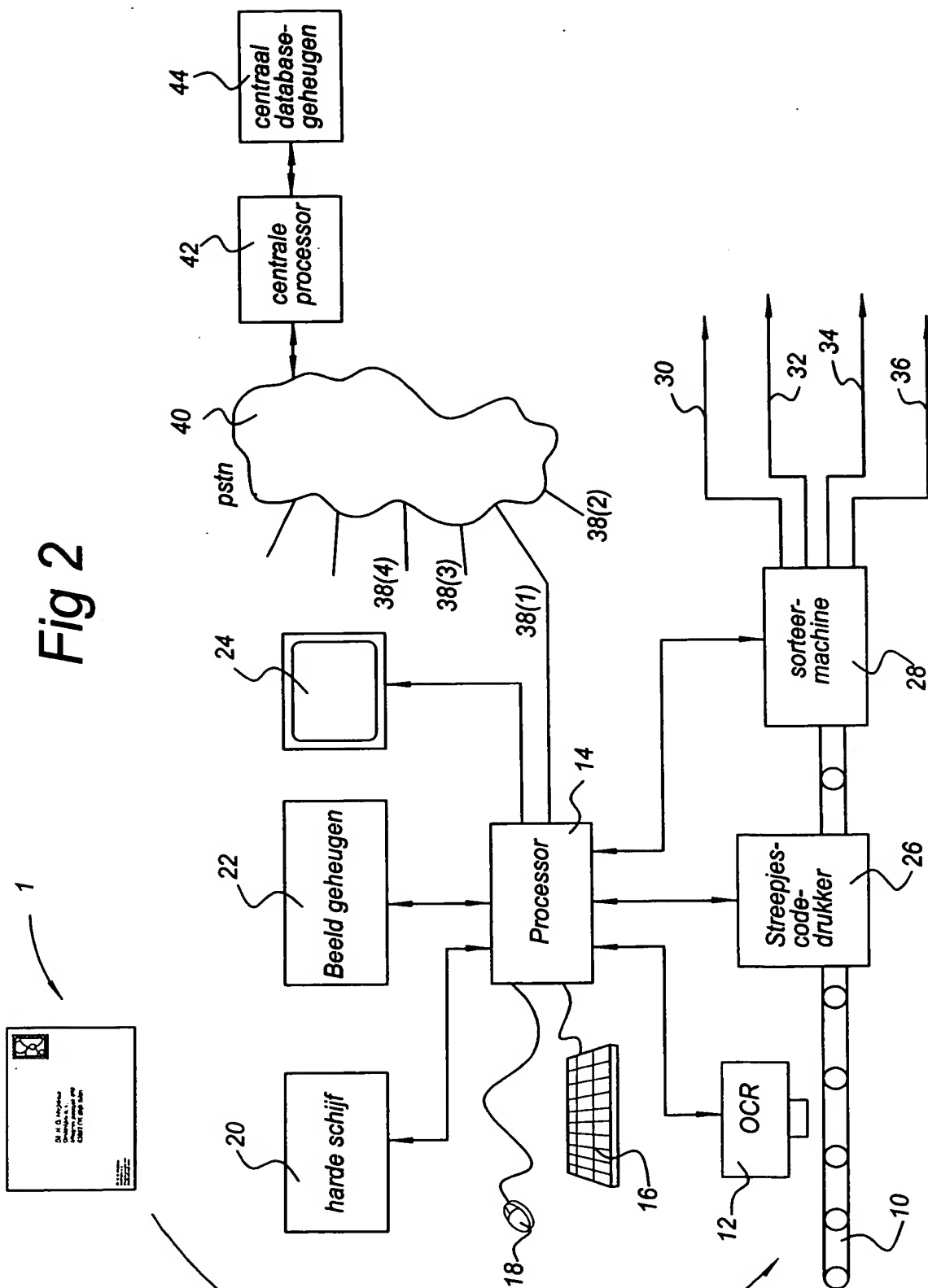
- 5 JULI 1999

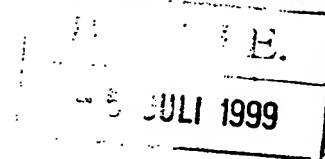
Uittreksel

Inrichting en werkwijze voor het actualiseren van een adressendatabase in een databasegeheugen (44) met geregistreerde adresrecords onder gebruikmaking van de volgende stappen:

- het inlezen van op poststukken (1) aanwezige adresgegevens (2);
- het verwerken van op poststukken vermelde adresgegevens;
- het opslaan van de adresgegevens;
- het vaststellen van een kwaliteitsmaat voor de adresgegevens op basis van vooraf bepaalde criteria;
- het vergelijken van de adresgegevens met de in het databasegeheugen (44) opgeslagen adresrecords; en
- het actualiseren van de inhoud van het databasegeheugen (44) op basis van de kwaliteitsmaat en het vergelijken van de adresgegevens met de opgeslagen adresrecords.

[fig. 2]





Inrichting en werkwijze voor het actualiseren van een adressendatabase met geregistreerde adresrecords

De onderhavige uitvinding heeft betrekking op een inrichting voor het actualiseren van een adressendatabase met geregistreerde adresrecords, omvattende:

- tenminste één processor voor het ontvangen en verwerken van op poststukken vermelde adresgegevens;
- een met de tenminste ene processor verbonden geheugen voor het opslaan van de adresgegevens;
- een met de tenminste ene processor verbonden databasegeheugen met daarin opgeslagen de adressendatabase.

Een dergelijke inrichting is bekend uit JP-A-9/57204.

JP-A-8/323299 heeft betrekking op een adresdatabasemanagementsysteem dat is ingericht voor het actualiseren van een adresdatabase op basis van teruggestuurde post waarop correcties zijn aangebracht. De teruggestuurde post wordt gelezen door een leeseenheid en gecorrigeerde gegevens worden in een geheugen opgeslagen. Het managementsysteem actualiseert automatisch de database. Een postsorteermachine is gekoppeld aan het databasemanagementsysteem dat post in overeenstemming met het gecorrigeerde adres kan sorteren. Dit document heeft geen betrekking op het actualiseren van bestaande databases zonder gebruik te maken van teruggestuurde post waarop correcties zijn aangebracht.

JP-A-9/314067 beschrijft een postsorteerinrichting en -werkwijze. Een opstelling wordt getoond die middelen omvat voor het uitlezen van adresinformatie uit de post en het op basis daarvan genereren van een adrescode. Als alternatief kan de adresinformatie met de hand worden ingevoerd. De adrescode wordt vergeleken met de inhoud van een adresinformatiedatabase. Wanneer de adrescode niet letterlijk correspondeert met een adres in de database, worden de adrescode en mogelijk correcte adressen aan de bediening getoond. De bediening kan dan, bijvoorbeeld door de naam van de geadresseerde op de post en de namen van de geadresseerde in de mogelijk correcte adressen beslissen welke daarvan correct is. Dit correcte adres wordt dan op het poststuk geprint voor verdere geautomatiseerde sortering. Dit document gaat uit van een correcte adresdatabase.

JP-A-9/75862 beschrijft een postsorteermachine. Deze omvat een videocodeerterminal. Bediening van de machine voert een gebruikerscode in. De machine evalueert de juistheid van de gebruikerscode via vergelijking met een database. Dit document maakt geen melding van actualiseren van een adresdatabase.

- 5 JP-A-9/57204, waarvan de conclusies zijn afgebakend, betreft het probleem van het actualiseren van een adresdatabase, die wordt gebruikt in een automatische postsorteerinrichting.

In een dergelijke postsorteerinrichting is een videocodeereenheid aanwezig, die poststukken uitleest en de adresgegevens bijvoorbeeld met een 'optical character recognition' eenheid omzet in een adrescode. De adrescode wordt opgeslagen in een geheugen. Tijdens het sorteerproces wordt de adrescode vergeleken met in een adresdatabase opgeslagen adresgegevens. Als uit de vergelijking blijkt, dat de adrescode correspondeert met een in de database opgeslagen adres, dan wordt dit voor de verdere sortering als correct aangenomen.

- 15 Sommige adrescodes corresponderen met geen enkel adres in de database. Het is dan mogelijk om het beeld dat door videocodeereenheid van het poststuk is geregistreerd op een scherm weer te geven. De bediening van de postsorteerinrichting bestudeert dan de op het scherm ingevoerde informatie en voert met de hand, bijvoorbeeld via een toetsenbord, het adres in dat door het verdere sorteerproces moet worden gebruikt. Het genoemde JP-A-9/57204 stelt echter ook voor om de database zelf te actualiseren, zodat fouten en/of gevallen waarin adrescodes niet corresponderen met een adres in de database steeds minder vaak zullen voorkomen.

Daartoe stelt JP-A-9/57204 in essentie het volgende voor. De adrescodes van de poststukken die niet corresponderen met een adres in de database worden in een geheugen opgeslagen. De bijbehorende, geweigerde poststukken worden via een aparte lijn afgevoerd. Op een geschikt tijdstip worden de geweigerde poststukken opnieuw de machine ingevoerd om op basis van hernieuwd uitlezen de database te actualiseren. Principieel kan de weigering van de poststukken twee verschillende oorzaken hebben. Allereerst kan het 'optical character recognition' systeem een of meer karakters niet hebben herkend. Ten tweede kan het zijn, dat alle karakters wel zijn herkend, maar dat de aldus vastgestelde adrescode niet correspondeert met een adres in de database.

In het eerste geval wordt het beeld dat door de videocodeereenheid is geregistreerd aan de bediening van het systeem getoond. De bediening beslist op basis daar-

van hoe de foutieve karakters moeten worden gecorrigeerd. Via patroonherkenning van het aldus met de hand gecorrigeerde adres wordt de database eventueel gecorrigeerd of aangevuld. In het tweede geval wordt de database aangevuld met de nieuwe adresinformatie, waarvan wordt aangenomen, dat deze correct is.

5 Het probleem dat door de onderhavige uitvinding wordt opgelost betreft het zo veel mogelijk automatisch actualiseren van een adressendatabase. Problemen die daarbij optreden zijn bijvoorbeeld als volgt. Veel mensen bevinden zich tijdelijk op een adres dat niet hun woonadres is, bijvoorbeeld als zij op vakantie zijn of als zij enige tijd in een ziekenhuis zijn opgenomen. In die tijd wordt post voor hen naar een ander adres
10 dan hun woonadres gestuurd. Desalniettemin is de adressering op de post voor hen correct en mag sortering niet worden geweigerd. Evenmin mag een adresdatabase dan worden aangepast. Bovendien verhuizen veel mensen zonder dat zij postbezorgingsbedrijven op tijd of op correcte wijze hun nieuwe woonadres meedelen. In dat geval is wijziging van de adresdatabase wel vereist. Ook worden veel mensen op
15 verschillende adressen geadresseerd, bijvoorbeeld omdat zij een privé-adres en een zakelijk adres hebben. Tenslotte is van sommige mensen (bijvoorbeeld kinderen) in het geheel geen adres bij postbezorgingsinstanties bekend, terwijl dat voor controle van het sorteerproces wel handig zou zijn.

 De doelstelling van de aanvraag is om bij het actualiseren van de
20 adressendatabase dit soort problemen zoveel mogelijk te vermijden en zo betrouwbaar mogelijke gegevens te verkrijgen.

 Om dit soort problemen op te lossen heeft de inrichting van de bij de aanvang genoemde soort het kenmerk, dat de tenminste ene processor is ingericht om voor de adresgegevens een kwaliteitsmaat vast te stellen op basis van vooraf bepaalde criteria,
25 de adresgegevens te vergelijken met de in het databasegeheugen opgeslagen adresrecords, en op basis van de kwaliteitsmaat en het vergelijken van de adresgegevens met de opgeslagen adresrecords de inhoud van het databasegeheugen te actualiseren.

 In de inrichting van de uitvinding wordt een adressendatabase met geregistreerde
30 adresrecords dus voortdurend geactualiseerd op basis van door de inrichting verwerkte poststukken. Het actualiseren van de adressendatabase gebeurt geheel automatisch en is niet, zoals in de stand van de techniek, gebaseerd op het handmatig verbeteren van de adressendatabase op basis van door de inrichting geweigerde poststukken, waarvan de

gelezen adresgegevens niet corresponderen met opgeslagen adresrecords. Daartoe maakt de inrichting volgens de uitvinding gebruik van een kwaliteitsmaat die aangeeft hoe goed de adresgegevens zijn en wordt vastgesteld op basis van vooraf bepaalde criteria. De inrichting zal geheel automatisch bepalen of de kwaliteit van die orde is dat
5 de adressendatabase kan worden geactualiseerd op basis van de adresgegevens.

Bij het actualiseren van de adressendatabase kan gebruik gemaakt worden van alle door de inrichting van poststukken ingelezen adresgegevens. Uiteraard is het echter ook mogelijk om een selectie daarbij toe te passen. Gedacht kan bijvoorbeeld worden aan aselechte steekproeven of het vermijden van het actualiseren van de adressendata-
10 base gedurende een bepaalde periode, waarin veel post wordt aangeboden voor ongebruikelijke adressen, bijvoorbeeld de kerstperiode. Ook is het mogelijk om bijvoorbeeld handmatig geschreven adresgegevens niet te verwerken, of adresgegevens van ansichtkaarten buiten beschouwing te laten.

De onderhavige uitvinding heeft ook betrekking op een werkwijze voor het actualiseren van een adressendatabase in een databasegeheugen met geregistreeerde adres-
15 records, omvattende:

- het inlezen van op poststukken aanwezige adresgegevens;
- het verwerken van op poststukken vermelde adresgegevens;
- het opslaan van de adresgegevens;

20 gekenmerkt door

het vaststellen van een kwaliteitsmaat voor de adresgegevens op basis van vooraf bepaalde criteria, het vergelijken van de adresgegevens met de in het databasegeheugen opgeslagen adresrecords, en het actualiseren van de inhoud van het databasegeheugen op basis van de kwaliteitsmaat en het vergelijken van de adresgegevens met de
25 opgeslagen adresrecords.

Tevens heeft de onderhavige uitvinding betrekking op een gegevensdrager voorzien van een computerprogramma dat door een computerinrichting kan worden gelezen en na te zijn ingelezen de computerinrichting de functionaliteit van het actualiseren van een adressendatabase in een databasegeheugen met geregistreeerde adresrecords ver-
30 schaft onder gebruikmaking van de volgende stappen:

- het inlezen van op poststukken aanwezige adresgegevens;
 - het verwerken van op poststukken vermelde adresgegevens;
 - het opslaan van de adresgegevens;
-

gekenmerkt door

het vaststellen van een kwaliteitsmaat voor de adresgegevens op basis van vooraf bepaalde criteria, het vergelijken van de adresgegevens met de in het databasegeheugen opgeslagen adresrecords, en het actualiseren van de inhoud van het databasegeheugen op basis van de kwaliteitsmaat en het vergelijken van de adresgegevens met de opgeslagen adresrecords.

Voorts heeft de onderhavige uitvinding betrekking op een computerprogramma dat door een computerinrichting kan worden gelezen en na te zijn ingelezen de computerinrichting de functionaliteit van het actualiseren van een adressendatabase in een databasegeheugen met geregistreerde adresrecords verschaft onder gebruikmaking van de volgende stappen:

- het inlezen van op poststukken aanwezige adresgegevens;
- het verwerken van op poststukken vermelde adresgegevens;
- het opslaan van de adresgegevens;

gekenmerkt door

het vaststellen van een kwaliteitsmaat voor de adresgegevens op basis van vooraf bepaalde criteria, het vergelijken van de adresgegevens met de in het databasegeheugen opgeslagen adresrecords, en het actualiseren van de inhoud van het databasegeheugen op basis van de kwaliteitsmaat en het vergelijken van de adresgegevens met de opgeslagen adresrecords.

De uitvinding zal hierna worden toegelicht onder verwijzing naar enkele figuren, die slechts zijn bedoeld ter illustratie van de uitvinding en niet ter beperking van de reikwijdte daarvan.

Figuur 1 is een voorbeeld van een poststuk dat is voorzien van adresgegevens van een geadresseerde en adresgegevens van een afzender;

figuur 2 is een voorbeeld van een inrichting die kan worden gebruikt bij het actualiseren van een database met adresrecords;

figuur 3 is een stroomschema van enkele stappen tijdens het proces van actualiseren van adresrecords in een database;

figuren 4, 5 en 6 geven nadere details van enkele van de in figuur 3 getoonde stappen weer.

In figuur 1 is een voorbeeld van een brief getoond, die is voorzien van adresgegevens van een geadresseerde 2 en adresgegevens van een afzender 7. De brief 1 kan

zijn voorzien van een postzegel 8. In plaats van een postzegel 8 kan ook een frankeerstempel of dergelijke zijn aangebracht.

De adresgegevens van de geadresseerde 2 omvatten bijvoorbeeld een naamregel 3, een firmanaam 4, een straatnaam 5 en een plaatsnaam met postcode 6. De adresgegevens van de afzender kunnen op soortgelijke wijze zijn opgebouwd (niet nader getoond in figuur 1).

Een dergelijke brief 1 kan met een automatische sorteermachine worden gesorteerd. Een voorbeeld daarvan is getoond in figuur 2. De brief 1 wordt bijvoorbeeld op een lopende band 10 gelegd. De lopende band 10 leidt de brief 1 langs een uitleeseenheid 12, die de adresgegevens van de geadresseerde 2, en bij voorkeur ook de adresgegevens van de afzender 7, kan uitlezen. De uitleeseenheid 12 is bijvoorbeeld een "optical character recognition"-eenheid, die aan de vakman genoegzaam bekend is.

Na te zijn ingelezen door de uitleeseenheid 12 voert de lopende band 10 de brief 1 naar een streepjescodedrukker 26, die een streepjescode op de brief 1 kan afdrukken. Deze streepjescode omvat nadere gegevens die bij de uiteindelijke sortering worden gebruikt. Na door de streepjescodedrukker 26 te zijn geleid, voert de lopende band 10 de brief 1 naar een sorteermachine 28, die de afgedrukte streepjescode inleest en op basis daarvan de brief 1 een bepaalde route 30, 32, 34, 36, ... inleidt.

De uitleeseenheid 12, de streepjescodedrukker 26 en de sorteermachine 28 zijn verbonden met een processor 14. De processor 14 is tevens verbonden met middelen waarmee de bediening gegevens kan invoeren, zoals een muis 18 en een toetsenbord 16. Ook is de processor 14 verbonden met geschikte geheugens, waaronder bijvoorbeeld een harde schijf 20 en een beeldgeheugen 22, waarin de processor 14 de door de uitleeseenheid 12 uitgelezen adresgegevens, zowel met betrekking tot de geadresseerde als de afzender (indien van toepassing) op kan slaan.

De processor 14 is tevens verbonden met weergeefmiddelen 24, bijvoorbeeld in de vorm van een monitor.

De processor 14 bestuurt de uitleeseenheid 12, de streepjescodedrukker 26 en de sorteermachine 28 ten behoeve van het sorteerproces op automatische wijze. De wijze waarop dit gebeurt is op zichzelf bekend en behoeft hier geen nadere toelichting.

De processor 14 is in de opstelling volgens figuur 2 verbonden met een telefoonnetwerk, bijvoorbeeld PSTN 40, (Public Switched Telephone Network).

Een centrale processor 42 is eveneens verbonden met PSTN 40. Tevens is de centrale processor 42 verbonden met een centraal databasegeheugen 44.

In figuur 2 is aangegeven dat de processor 14 met de PSTN 40 is verbonden via een kabel 38(1) en dat er meerdere van dergelijke verbinding 38(2), 38(3), 38(4), ...
 5 zijn. Deze overige verbindingen zijn bedoeld om aan te geven dat er meerdere van dergelijke opstellingen met lokale processoren en lokale sorteermiddelen op het PSTN 40 kunnen zijn aangesloten. Zij kunnen alle via het PSTN 40 met de centrale processor 42 communiceren.

Het is echter ook denkbaar dat de uitvinding wordt toegepast op een lokale op-
 10 stelling en dat de processor 14 direct verbonden is met het centrale databasegeheugen 44.

Het centrale databasegeheugen omvat een database met adresrecords. Een adres-
 record bevat tenminste een naamregel 3, een straatnaam 5 en een plaatsnaam met post-
 code 6. Indien van toepassing kan ook een bedrijfsnaam 4 zijn opgenomen. Uiteraard is
 15 het ook mogelijk om verdere relevante gegevens in het adresrecord op te nemen.

Figuur 3 geeft in grote lijnen de stappen weer die worden gezet tijdens het actua-
 liseren van adresrecords in het centrale databasegeheugen 44.

Op dit moment worden in Nederland door middel van een Automatisch PostSor-
 teerSysteem (APSS) op zes verschillende sorteercentra dagelijks ongeveer 24.000.000
 20 poststukken automatisch verwerkt. Een zo'n sorteercentrum komt qua functionaliteit
 overeen met het middengedeelte van figuur 2. De uitleeseenheid 12 maakt een digitaal
 beeld van de adreszijde van elk poststuk 1. Het adres in het digitale beeld wordt gelo-
 kaliseerd en vervolgens wordt het adres automatisch gelezen. In een eerste stap van het
 stroomdiagram in figuur 3, "APSS-collectie" 46, worden de adresgegevens van de post-
 25 stukken tijdelijk lokaal opgeslagen in een beeldgeheugen 22. Bij voorkeur worden al-
 leen die adresgegevens gebruikt die in machineschrift op het poststuk 1 zijn aange-
 bracht en waarvan de postcode/huisnummer-combinatie 6 correct is gelezen. De adres-
 gegevens van overige post worden bij voorkeur als onvoldoende betrouwbaar be-
 schouwd om de adresdatabase in het centrale databasegeheugen 44 te actualiseren.

30 De centrale processor 42 kan via de netwerkverbinding PSTN 40 (of via elke
 andere geschikt communicatieverbinding) communiceren met processor 14 om de in
 het beeldgeheugen 22 opgeslagen beelden op te vragen. Na ontvangst kan de centrale
 processor 42 eventueel noodzakelijke administratieve gegevens aan de ingelezen adres-

gegevens toevoegen. Administratieve gegevens bestaan bijvoorbeeld uit het sorteercentrum waar de adresgegevens zijn ingelezen, de datum van uitlezing van de adresgegevens en een volgnummer. Daarna kan de centrale processor 42 de ontvangen, ingelezen adresgegevens centraal opslaan in elk geschikt geheugen. Dit kan bijvoorbeeld
 5 een geheugen zijn dat apart is van het centrale databasegeheugen 44.

Na de "APSS-collectie"-stap 46 wordt een segmentatie/normalisatie-stap 48 uitgevoerd. Zoals getoond in figuur 4 worden in deze stap de opgehaalde, ingelezen adresgegevens (stap 482) per poststuk in relevante onderdelen gesplitst. Uit de adresgegevens 2 van de geadresseerde worden naamregels 3 geselecteerd, stap 483, en
 10 opgesplitst in delen, stap 484. Aan de afzonderlijke delen worden betekenissen toegekend, zoals voorvoegsel, voornaam, titel, etc. Vervolgens worden de afzonderlijke delen genormaliseerd. Indien aanwezig wordt ook de firmanaam 4 geanalyseerd. Er resulteert dan een adresrecord met achternamen en/of bedrijfsnamen en bijbehorende attributen, zoals voorletters, tussenvoegsels, achtervoegsels, etc. Volgens een vooraf
 15 bepaald algoritme wordt uit de gegevens in het adresrecord een algehele herkenbetrouwbaarheid bepaald, stap 485. Ook wordt een maat voor de kwaliteit van de segmentatie (het in afzonderlijke delen splitsen) binnen het segmentatieproces bepaald, stap 486. Zowel de herkenbetrouwbaarheid als de segmentatiekwaliteit worden opgenomen in het adresrecord. Beide maten, dat wil zeggen
 20 herkenbetrouwbaarheid en segmentatiekwaliteit, worden later gebruikt bij het daadwerkelijk actualiseren van de adresrecords in het centrale databasegeheugen 44.

Vervolgens wordt een gangbaarheidstest 50 door de centrale processor 42 uitgevoerd (zie figuur 5). Daartoe beschikt de centrale processor 42 over een bestand van in Nederland voorkomende achternamen. Dit bestand is in een (niet weergegeven) met de
 25 centrale processor 42 verbonden geheugen opgeslagen. Elke achternaam die uit de adresgegevens is gesplitst met behulp van stap 48 wordt in stap 501 vergeleken met voorkomende achternamen en in stap 502 aan syntactische regels getoetst. Hetzelfde gebeurt met voornamen en eventueel aanwezige bedrijfsnamen. Met behulp van een vooraf bepaald algoritme wordt een gangbaarheidsmaat van elke achternaam,
 30 voornaam en eventuele bedrijfsnaam bepaald en aan het adresrecord toegevoegd, stap 503.

Na het uitlezen met behulp van uitleeseenheid 12 zijn de straatnaamgegevens 5 (of postbusnummers) en de combinatie van postcode en plaatsnaam 6 reeds door de

processor 14 getest op de bestaanbaarheid. In de huidige praktijk worden huisnummertoevoegsels, zoals "boven", "achter", "2-hoog", "II", enz. niet gecontroleerd. In stap 52 "adressering" worden deze huisnummertoevoegsels genormaliseerd aan de hand van een lijst met veel voorkomende afkortingen. "TO" is bijvoorbeeld synoniem met
 5 "tegenover". "1-hoog" is bijvoorbeeld synoniem met "1^e" en met "i".

In stap 52 controleert de centrale processor 42 of er voor elke combinatie van straatnaam, huisnummer, postcode en plaats corresponderende gegevens in de, in het centrale databasegeheugen 44 opgeslagen database aanwezig zijn. Is dit niet het geval, dan is het uitgelezen adres "nieuw". Dit gegeven wordt aan het ingelezen adresrecord
 10 toegevoegd (stap 522).

De centrale processor 42 leest vervolgens alle op het adres (inclusief toevoegsels) van een ingelezen adresrecord ingeschreven namen uit de centrale database in het centrale databasegeheugen 44. De ingelezen namen 3, 4 worden vergeleken met de ingeschreven namen uit de centrale database. Indien mogelijk worden eventuele voorletters,
 15 voornamen, voorvoegsels, achternamen en achtervoegsels met elkaar vergeleken. Elke ingelezen naam, 3, 4 van een geadresseerde krijgt een vergelijkingsscore, waarin de kwaliteit en de resultaten van vergelijkingen met de verschillende componenten van de ingeschreven naam, te weten voorletters, voorvoegsels, voornaam, achternaam, achtervoegsels, enz., zijn opgenomen.

Op basis van de vergelijkingsscores, de segmentatiekwaliteit, de gangbaarheidsmaat en de herkenbetrouwbaarheid van de ingelezen adresgegevens besluit de centrale processor 42 of een ingelezen adresrecord nieuw, bekend of onbekend is. Dat kan bijvoorbeeld in overeenstemming met de volgende regels:

- de ingelezen adresgegevens zijn nieuw als de vergelijkingsscores een relatief geringe waarde hebben en de kwaliteitsmaat relatief groot is;
- de ingelezen adresgegevens zijn bekend als de vergelijkingsscores een relatief grote waarde hebben en de kwaliteitsmaat relatief groot is;
- de ingelezen adresgegevens zijn onbekend als de vergelijkingsscores een relatief geringe waarde hebben en de kwaliteitsmaat relatief gering is.

30 Dit gebeurt in stap 56 "besluiten".

In dezelfde stap wordt de statistiek van bekende adresrecords bijgewerkt. Hierop wordt hieronder nog nader ingegaan.

Voor een nieuw en/of onbekend adresrecord kan de centrale processor 42 een nieuw adresrecord in het centrale databasegeheugen 44 genereren.

Voor het bijwerken van statistische gegevens met betrekking tot reeds bekende adresrecords kunnen bijvoorbeeld de volgende stappen worden gezet.

5 De frequentie van voorkomen van een gegeven kan per sorteercentrum worden bijgehouden, zodat de spreiding van een adresgegeven kan worden vastgelegd. Ook kan worden bijgehouden op welke datum een adresrecord voor het laatst in een van de sorteercentra werd ingelezen. Aldus kan in de centrale database worden bijgehouden op welke data van bijvoorbeeld het laatste jaar een betreffend adresrecord ergens op een
10 poststuk is voorgekomen. Dan kan het interval tussen twee opeenvolgende keren dat het adresrecord op een poststuk werd gebruikt, worden bepaald. Voor het bepalen van dat interval kan bijvoorbeeld worden gekeken naar een recente periode of naar een langer geleden periode. Ook kan het gemiddelde worden bepaald tussen twee opeenvolgende keren dat het adresrecord in een van de sorteercentra werd uitgelezen. Indien het
15 gemiddelde in de loop van de tijd sterk wijzigt, kan dit een aanwijzing zijn dat het adresrecord moet worden gewijzigd.

Het is ook mogelijk dat de adresgegevens 7 van de afzender worden geregistreerd en in een geheugen worden opgeslagen. Vervolgens kunnen de afzendergegevens 7 dan worden gebruikt voor het toekennen van een kwaliteitsmaat. Er kan immers van wor-
20 den uitgegaan, dat poststukken die afkomstig zijn van bijvoorbeeld de burgerlijke stand van een gemeente zeer betrouwbare adresgegevens van de geadresseerde 2 bevatten.

Op basis van de gegevens kan elk adresrecord een status worden toegekend, bijvoorbeeld nieuw, gangbaar, betrouwbaar, of oud. Daartoe kunnen bijvoorbeeld de volgende regels gelden. Wanneer een adresrecord voor het eerst wordt aangemaakt, krijgt
25 het de status nieuw. Een adresrecord wordt gangbaar wanneer het meerdere keren van verschillende afzenders en/of van verschillende sorteercentra wordt ontvangen. Een gangbaar adresrecord kan betrouwbaar worden, wanneer het adresrecord regelmatig op poststukken wordt aangetroffen. De regelmaat wordt bepaald aan de hand van het gemiddelde interval tussen twee opeenvolgende keren dat het adresrecord op een poststuk
30 wordt ingelezen. Indien de gemiddelde duur van dat interval langer wordt dan een vooraf bepaalde tijdsduur kan worden bepaald, dat een betrouwbaar adresrecord weer gangbaar wordt. Indien een adresrecord vervalt, krijgt het de status oud.

Indien het interval tussen recente keren dat een adresrecord op een poststuk wordt aangetroffen steeds langer wordt, is dit een indicatie van bijvoorbeeld verhuizing en kan worden besloten dat het adresrecord moet vervallen.

In figuur 2 is de situatie getekend, dat het centrale databasegeheugen één database heeft. Het is echter ook mogelijk om op basis van de daarin aanwezige database een extra database te maken, dat een selectie van de adresrecords uit de gehele centrale database omvat. Deze selectie vindt plaats op basis van bepaalde beslisregels. In deze selectie van adresrecords zijn bijvoorbeeld alleen de betrouwbare adresrecords opgenomen. Deze betrouwbare adresrecords kunnen dan bijvoorbeeld beschikbaar worden gesteld voor gebruik door derden.

In het geval dat is voorzien in een extra database met een selectie van de gegevens uit het centrale databasegeheugen, kan er in zijn voorzien dat als extra beslisregel voor het opnemen in het extra databasegeheugen geldt, dat selectie in overeenstemming moet zijn met de betreffende regels van privacy-wetgeving.

Bij de besluitvorming over de bruikbaarheid van een ingelezen adresgegeven wordt dus bij voorkeur gebruik gemaakt van (1) de frequentie van voorkomen daarvan, (2) de variatie daarin en (3) de spreiding met betrekking tot afzenders. Daarvoor kan in principe een regel-gebaseerd systeem worden opgezet. Als alternatief kan de centrale processor 42 worden voorzien van een neurale netwerk of dergelijke. In het algemeen zal gelden dat een adresgegeven beter bruikbaar is naarmate het meer recent op een poststuk is uitgelezen, als het voldoende vaak voorkomt (gangbaar is), naarmate het van meer verschillende afzenders wordt ontvangen (valide) en als de kwaliteit daarvan, bijvoorbeeld aangeduid door de segmentatiekwaliteit en de herkenbetrouwbaarheid groter is.

Hierboven is een uitvoeringsvorm beschreven waarin de stappen voor het actualiseren van de centrale database worden uitgevoerd door de centrale processor 42. Zoals eerder opgemerkt, kunnen deze stappen echter ook worden uitgevoerd door de processor 14, namelijk dan wanneer het gaat om het actualiseren van een adresdatabase dat alleen lokaal voorkomt en dat is opgeslagen in een geheugen dat met de processor 14 is verbonden.

Extra mogelijkheden voor de beschreven inrichting en werkwijze zijn als volgt.

Bij het uitlezen met de uitleeseenheid 12 kunnen de adresgegevens van de afzender 7 worden ingelezen. De processor 14 kan bijvoorbeeld vaststellen dat de adres-

- gegevens van de afzender 7 overeenkomen met die van de PTT POST (of het sorteerbedrijf) zelf. De daarvan afkomstige gegevens zullen corresponderen met de gegevens in het centrale databasegeheugen 44. Dergelijke gegevens kunnen bij voorbaat niet worden opgenomen in het beeldgeheugen 22 om te voorkomen dat eventueel fouten in
- 5 het centrale databasegeheugen 44 zichzelf iedere keer bevestigen.

Behalve titels en dergelijke van namen in naamregels 3 komt het voor, dat functies van mensen op poststukken staan vermeld. Ook deze gegevens kunnen in het centrale databasegeheugen 44 worden opgeslagen. Deze functies kunnen van belang zijn bij het opbouwen van de genoemde extra database.

Conclusies

1. Inrichting voor het actualiseren van een adressendatabase met geregistreerde
 5 adresrecords, omvattende:
 - een leeseenheid (12) voor het inlezen van op poststukken (1) aanwezige adres-
 gegevens (2);
 - tenminste één processor (14, 42) voor het ontvangen en verwerken van op
 poststukken vermelde adresgegevens;
 - 10 • een met de tenminste ene processor (14, 42) verbonden geheugen (22) voor het
 opslaan van de adresgegevens;
 - een met de tenminste ene processor (14, 42) verbonden databasegeheugen (44)
 met daarin opgeslagen de adressendatabase;
 met het kenmerk, dat
- 15 de tenminste ene processor (14, 42) is ingericht om voor de adresgegevens een
 kwaliteitsmaat vast te stellen op basis van vooraf bepaalde criteria, de adresgegevens te
 vergelijken met de in het databasegeheugen (44) opgeslagen adresrecords, en op basis
 van de kwaliteitsmaat en het vergelijken van de adresgegevens met de opgeslagen
 adresrecords de inhoud van het databasegeheugen (44) te actualiseren.
- 20 2. Inrichting volgens conclusie 1, waarbij de tenminste ene processor is ingericht
 om een betrouwbaarheidsmaat van herkenning van de adresgegevens te bepalen en de
 kwaliteitsmaat mede te baseren op de betrouwbaarheidsmaat van herkenning.
- 25 3. Inrichting volgens conclusie 1 of 2, waarbij de tenminste ene processor is inge-
 richt om uit de adresgegevens naamregels te selecteren, de naamregels volgens vooraf
 bepaalde regels in afzonderlijke delen te splitsen en de kwaliteitsmaat mede te baseren
 op het selecteren van naamregels en het splitsen daarvan.
- 30 4. Inrichting volgens conclusie 3, verder omvattende opgeslagen gangbare namen,
 en waarbij de tenminste ene processor is ingericht om de afzonderlijke delen van de
 naamregels te vergelijken met de gangbare namen, op basis daarvan een gangbaar-

heidsmaat vast te stellen en de kwaliteitsmaat mede te baseren op de gangbaarheidsmaat.

5. Inrichting volgens conclusie 3 of 4, waarbij de tenminste ene processor is ingericht om uit de naamregels een naam van een geadresseerde af te leiden, uit de adresgegevens een adres van de geadresseerde af te leiden, ingeschreven namen van bij dat adres horende personen uit de adressendatabase uit te lezen en te vergelijken met de naam van de geadresseerde en op basis van die vergelijking een vergelijkingsscore per ingeschreven naam te bepalen, waarbij een vergelijkingsscore een grotere waarde heeft naarmate er meer overeenstemming is tussen de naam van de geadresseerde en een respectieve ingeschreven naam.

6. Inrichting volgens conclusie 5, waarbij de tenminste ene processor is ingericht om te bepalen, dat de adresgegevens

- nieuw zijn, als de vergelijkingsscores een relatief geringe waarde hebben én de kwaliteitsmaat relatief groot is;
- bekend zijn, als de vergelijkingsscores een relatief grote waarde hebben én de kwaliteitsmaat relatief groot is;
- onbekend zijn, als de vergelijkingsscores een relatief geringe waarde hebben én de kwaliteitsmaat relatief gering is.

7. Inrichting volgens conclusie 6, waarbij de tenminste ene processor is ingericht om als de adresgegevens nieuw zijn een extra, de adresgegevens bevattend adresrecord in de adressendatabase te genereren.

8. Inrichting volgens conclusie 7, waarbij de tenminste ene processor is ingericht om per adresrecord één van de volgende vier statussen te registreren:

- status nieuw, wanneer het adresrecord wordt gegenereerd;
- status gangbaar, wanneer de bijbehorende adresgegevens van verschillende afzenders worden ontvangen;
- status betrouwbaar, als de bijbehorende adresgegevens regelmatig opnieuw worden ingelezen;
- status oud, als het adresrecord verval.

9. Inrichting volgens een van de voorgaande conclusies, verder omvattende post-sorteereenheden (26, 28) voor het automatisch sorteren van de poststukken (1) onder gebruikmaking van de adressendatabase.
- 5 10. Werkwijze voor het actualiseren van een adressendatabase in een databasegeheugen (44) met geregistreerde adresrecords, omvattende:
- het inlezen van op poststukken (1) aanwezige adresgegevens (2);
 - het verwerken van op poststukken vermelde adresgegevens;
 - het opslaan van de adresgegevens;
- 10 gekenmerkt door
- het vaststellen van een kwaliteitsmaat voor de adresgegevens op basis van vooraf bepaalde criteria, het vergelijken van de adresgegevens met de in het databasegeheugen (44) opgeslagen adresrecords, en het actualiseren van de inhoud van het databasegeheugen (44) op basis van de kwaliteitsmaat en het vergelijken van de
- 15 adresgegevens met de opgeslagen adresrecords.
11. Werkwijze volgens conclusie 10, omvattende de stap van het bepalen van een betrouwbaarheidsmaat van herkenning van de adresgegevens en het mede baseren van de kwaliteitsmaat op de betrouwbaarheidsmaat van herkenning.
- 20 12. Werkwijze volgens conclusie 10 of 11, omvattende de stappen van het selecteren van naamregels uit de adresgegevens, het splitsen van de naamregels volgens vooraf bepaalde regels in afzonderlijke delen en het mede baseren van de kwaliteitsmaat op het selecteren van naamregels en het splitsen daarvan.
- 25 13. Werkwijze volgens conclusie 12, omvattende de stappen van het vergelijken van de afzonderlijke delen van de naamregels met gangbare namen, het vaststellen van een gangbaarheidsmaat op basis daarvan en het mede baseren van de kwaliteitsmaat op de gangbaarheidsmaat.
- 30 14. Werkwijze volgens conclusie 12 of 13, omvattende de stappen van het afleiden uit de naamregels van een naam van een geadresseerde, het afleiden uit de adresgegevens van een adres van de geadresseerde, het uitlezen van ingeschreven na-
-

men van bij dat adres horende personen uit de adressendatabase en het vergelijken daarvan met de naam van de geadresseerde en het bepalen van een vergelijkingsscore per ingeschreven naam op basis van die vergelijking, waarbij een vergelijkingsscore een grotere waarde heeft naarmate er meer overeenstemming is tussen de naam van de geadresseerde en een respectieve ingeschreven naam.

15. Werkwijze volgens conclusie 14, omvattende de stap van het bepalen, dat de adresgegevens
- nieuw zijn, als de vergelijkingsscores een relatief geringe waarde hebben én de kwaliteitsmaat relatief groot is;
 - bekend zijn, als de vergelijkingsscores een relatief grote waarde hebben én de kwaliteitsmaat relatief groot is;
 - onbekend zijn, als de vergelijkingsscores een relatief geringe waarde hebben én de kwaliteitsmaat relatief gering is.

15

16. Werkwijze volgens conclusie 15, omvattende de stap van het genereren van een extra, de adresgegevens bevattend adresrecord in de adressendatabase als de adresgegevens nieuw zijn.

17. Werkwijze volgens conclusie 16, omvattende de stap van het per adresrecord registreren van één van de volgende vier statussen:
- status nieuw, wanneer het adresrecord wordt gegenereerd;
 - status gangbaar, wanneer de bijbehorende adresgegevens van verschillende afzenders worden ontvangen;
 - status betrouwbaar, als de bijbehorende adresgegevens regelmatig opnieuw worden ingelezen;
 - status oud, als het adresrecord vervalt.

25

18. Werkwijze volgens een van de conclusies 10-17, omvattende de stap van het sorteren van de poststukken (1) onder gebruikmaking van de adressendatabase.

30

19. Gegevensdrager voorzien van een computerprogramma dat door een computerinrichting kan worden gelezen en na te zijn ingelezen de computerinrichting de functi-

onaliteit van het actualiseren van een adressendatabase in een databasegeheugen (44) met geregistreeerde adresrecords verschaft onder gebruikmaking van de volgende stappen:

- het inlezen van op poststukken (1) aanwezige adresgegevens (2);
- 5 • het verwerken van op poststukken vermelde adresgegevens;
- het opslaan van de adresgegevens;

gekenmerkt door

het vaststellen van een kwaliteitsmaat voor de adresgegevens op basis van vooraf bepaalde criteria, het vergelijken van de adresgegevens met de in het databasegeheugen
 10 (44) opgeslagen adresrecords, en het actualiseren van de inhoud van het databasegeheugen (44) op basis van de kwaliteitsmaat en het vergelijken van de adresgegevens met de opgeslagen adresrecords.

20. Computerprogramma dat door een computerinrichting kan worden gelezen en na
 15 te zijn ingelezen de computerinrichting de functionaliteit van het actualiseren van een adressendatabase in een databasegeheugen (44) met geregistreeerde adresrecords verschaft onder gebruikmaking van de volgende stappen:

- het inlezen van op poststukken (1) aanwezige adresgegevens (2);
- het verwerken van op poststukken vermelde adresgegevens;
- 20 • het opslaan van de adresgegevens;

gekenmerkt door

het vaststellen van een kwaliteitsmaat voor de adresgegevens op basis van vooraf bepaalde criteria, het vergelijken van de adresgegevens met de in het databasegeheugen
 (44) opgeslagen adresrecords, en het actualiseren van de inhoud van het
 25 databasegeheugen (44) op basis van de kwaliteitsmaat en het vergelijken van de adresgegevens met de opgeslagen adresrecords.

Fig 1

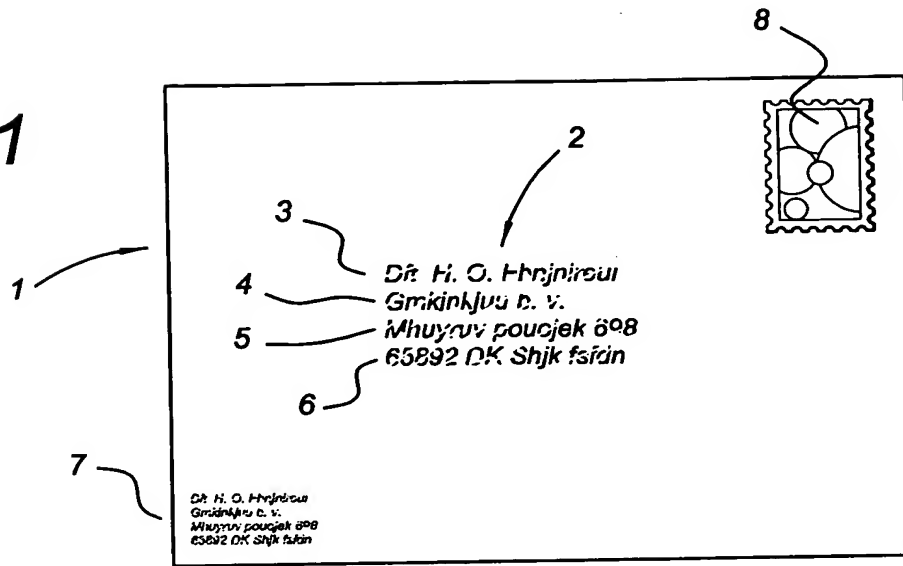
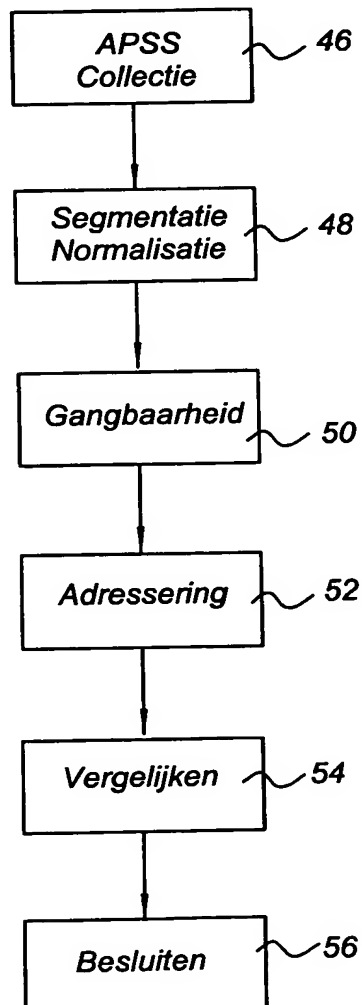


Fig 3



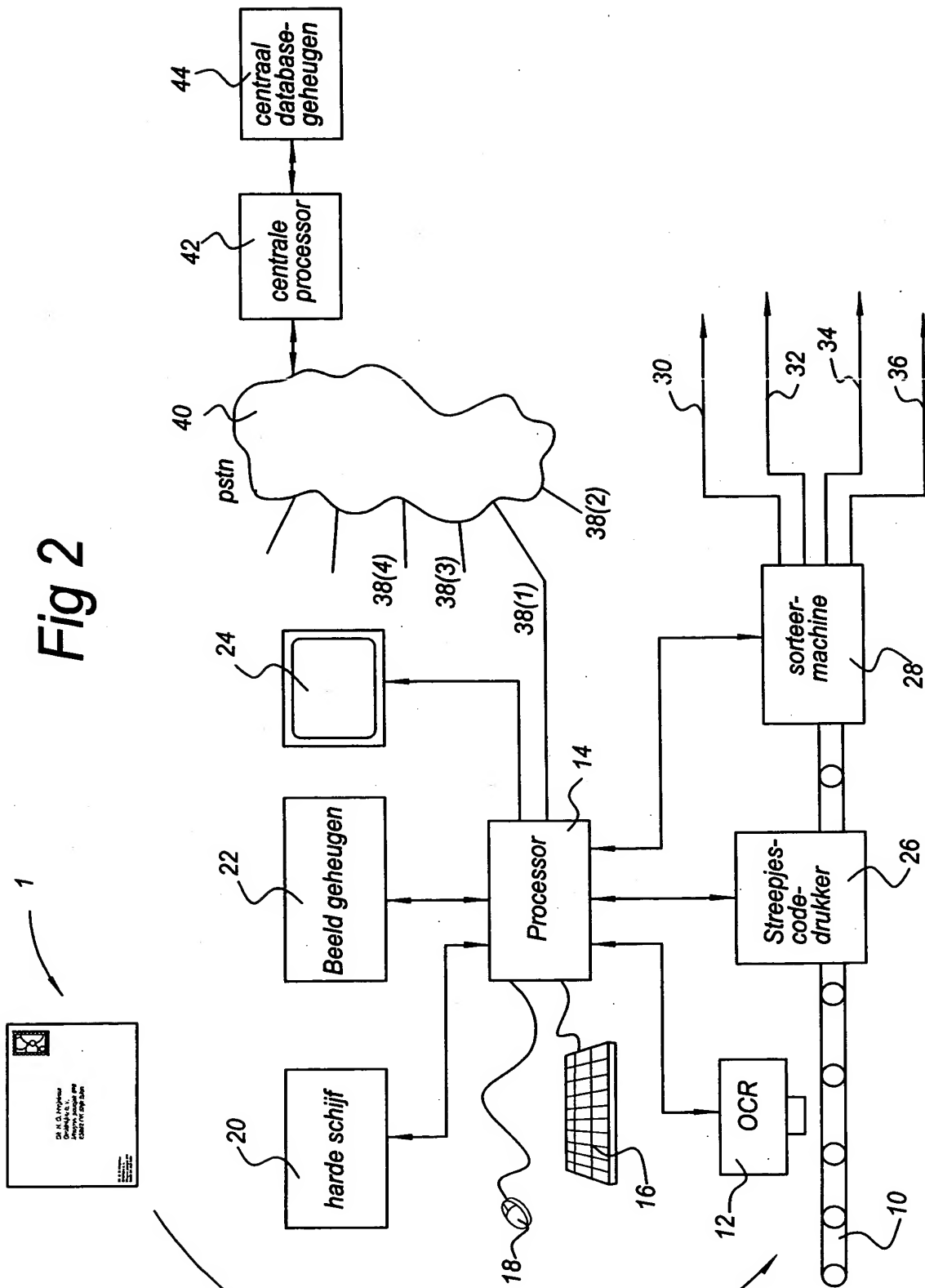


Fig 4

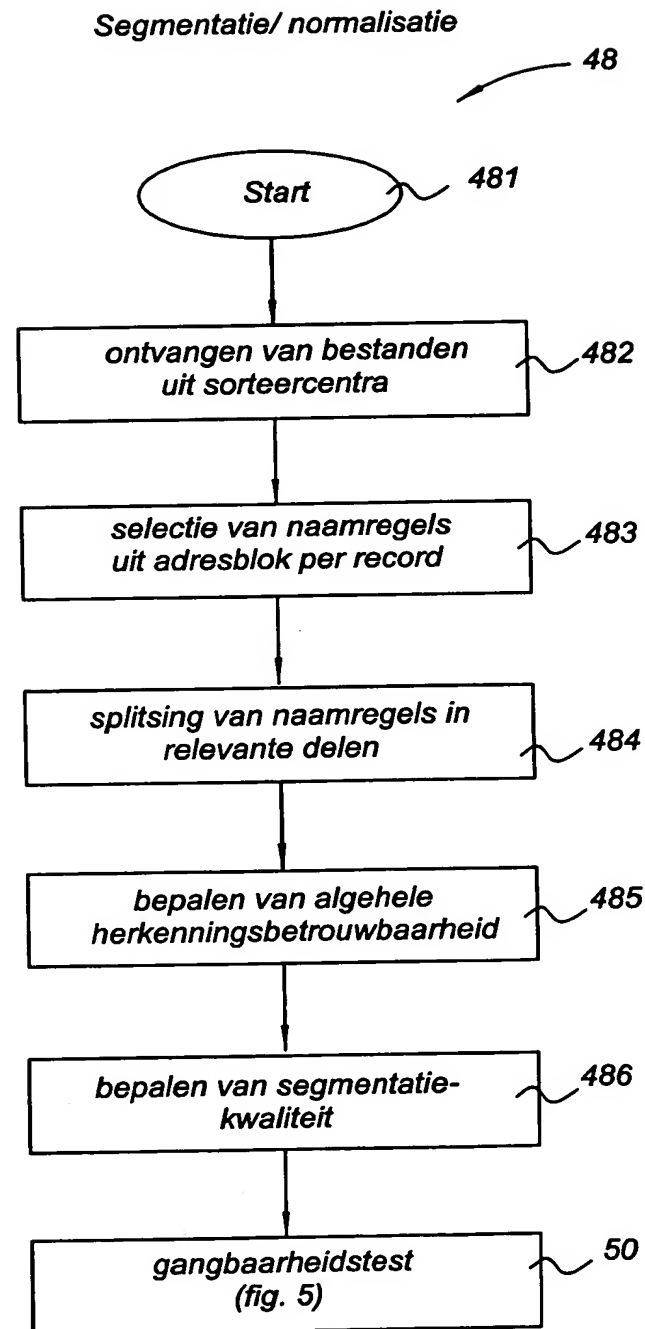


Fig 5

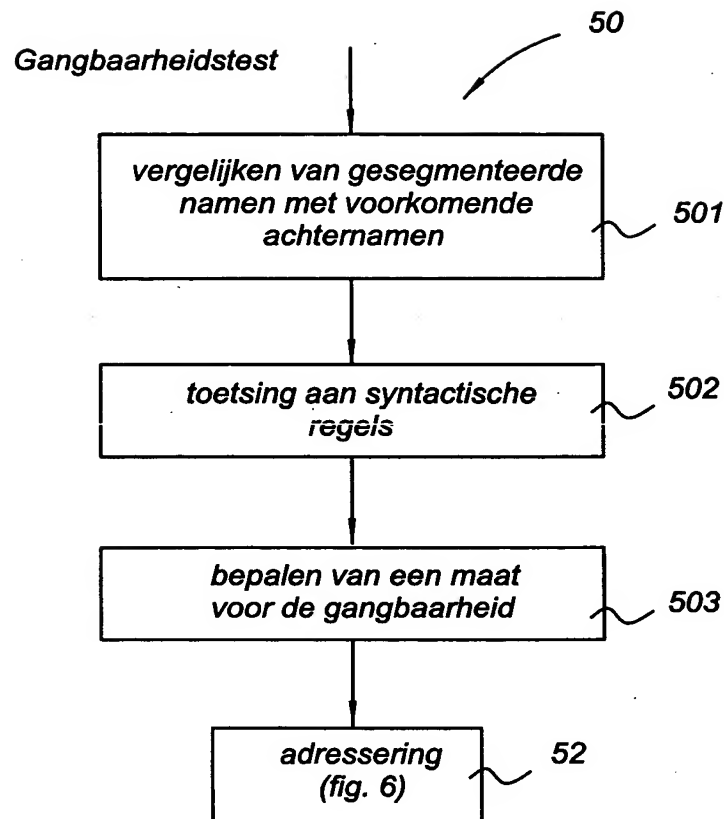


Fig 6

